IMERG on the Verge of Version 06

George J. Huffman[1], David T. Bolvin[1,2], Dan Braithwaite[3], Kuolin Hsu[3], Robert Joyce[4,5], Christopher Kidd[1,6], Eric Nelkin[1,2], Soroosh Sorooshian[3], Jackson Tan[1,7], Pingping Xie[5]

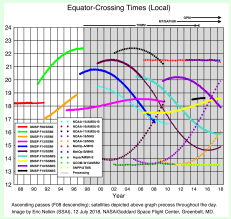
(1) NASA Goddard Space Flight Center
(2) Science Systems and Applications, Inc.
(3) Univ. of California Irvine
(4) Innovim
(5) NOAA/NWS Climate Prediction Center
(6) Univ. of Maryland / ESSIC
(7) Universities Space Research Assoc.

THE CURRENT GPM MICROWAVE CONSTELLATION

We want 3-hourly observations, globally
- Sampling the diurnal cycle
- Morphed microwave loses skill outside ±90 min

The current IMERG constellation includes:
- 5 polar-orbit passive microwave imagers
  - 3 SSMIS, AMSR-2, GMI
- 6 polar-orbit passive microwave sounders
  - 4 MHS, 2 ATMS (SAPHIR not yet contributing)
- Input **precip** estimates
  - GPROF (LEO PMW)
  - PERSIANN-CCS (GEO infrared)
  - 2BCMB (combined PMW-radar)
  - GPCP SG (monthly satellite-gauge)

The future is "interesting"
- Legacy satellites are allowed to drift
  - exact coverage is a complicated function of time
  - duplicate orbits aren't very useful for getting 3-hourly observations
- GPM fuel will last >10 years, so likely not the limiting factor
- Future launch manifests are assured for sounders, sparse for imagers
  - Microwave Imager (MWI) series – EUMETSAT
  - Weather System Follow-on-Microwave (WSF-M) series – DoD
    - perhaps at 0535 ECT descending
    - perhaps launching in 2022
  - Global Change Observation Mission-Water 3 (GCOM-W3) – JAXA (under consideration)



Ascending passes (FOR descending); satellites depicted above graph process throughout the day. Image by Eric Nelkin (SSAI), 12 July 2018, NASA/Goddard Space Flight Center, Greenbelt, MD.

IMERG DATASET CHARACTERISTICS

IMERG is a unified U.S. algorithm that takes advantage of the strengths of the partner algorithms
- Kalman Filter CMORPH – NOAA/CPC
- PERSIANN with Cloud Classification System – U.C.-Irvine
- TMPA – GSFC
- Precipitation Processing System (PPS, GSFC) computational environment

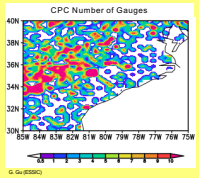
IMERG is a single integrated code system appropriate for near-real and post-real time
- Multiple runs accommodate different user requirements for latency and accuracy
  - "Early" – 4 hr (flash flooding)
  - "Late" – 14 hr (crop forecasting)
  - "Final" – 3 months (research)
- Time intervals are half-hourly and monthly (Final only)
- 0.1° global CED grid
  - morphed precip 60° N-S in V05, global in V06
  - IR covers 60° N-S
- User-oriented services by archive sites
  - interactive analysis (Giovanni)
  - alternate formats (TIFF files, …)
  - value-added products

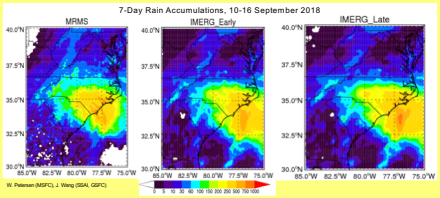
IMERG is adjusted to GPCP monthly climatology zonally to achieve a reasonable bias
- Over Version 04, 05, 06 the GPM core products have similar zonal profiles (by design)
  - these profiles are low in the extratropics compared to
    - GPCP monthly Satellite-Gauge product
    - Behrangi Multi-satellite CloudSat, TRMM, Aqua (MCTA) product
- Over land this provides a first cut at the adjustment to gauges that the final calibration in IMERG enforces

	Half-hourly data file (Early, Late, Final)
1	*[multi-sat.] precipitationCal*
2	[multi-sat.] precipitationUncal
3	*[multi-sat. precip] randomError*
4	[PMW] HQprecipitation
5	[PMW] HQprecipSource [identifier]
6	[PMW] HQobservationTime
7	IRprecipitation
8	IRkalmanFilterWeight
9	*[phase] probabilityLiquidPrecipitation*
10	*precipitationQualityIndex*
	Monthly data file (Final)
1	*[sat.-gauge] precipitation*
2	*[sat.-gauge precip] randomError*
3	GaugeRelativeWeighting
4	*probabilityLiquidPrecipitation [phase]*
5	*precipitationQualityIndex*

V05 VALIDATION – HURRICANE FLORENCE

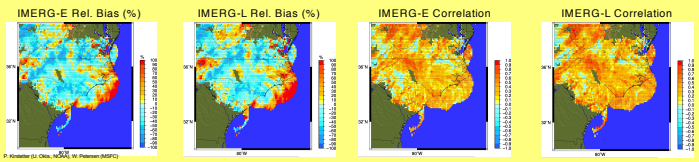
Florence approached the Carolina coast as Category 5 in early September, but then weakened to Category 1. Nonetheless, the forecast of extreme rain totals and extended flooding was accurate.
- Multi-Radar Multi-Sensor (MRMS) considered the best estimate
  - some questions about the details of the gauge calibration of the radar estimate
    - availability (gauge population figure to the right)
    - accuracy at high rates
  - limited to land and near-coastal areas





The overall appearance of the IMERG runs is similar to MRMS over land and near-shore waters (above)
- Looking more closely, the IMERG runs are closer to each other than to MRMS
  - Recall that the major difference is forward-only morphing in Early, but both forward and backward in Late
- The relative bias (below) shows that the Late Run is somewhat better
  - for both, there seems to be a shift north of the pattern along the coast (high in northeastern North Carolina and low along the Carolinas' border)
  - for both, there is underestimation along the eastern slope of the Appalachians
    - we suspect orographic enhancement not caught in GPROF

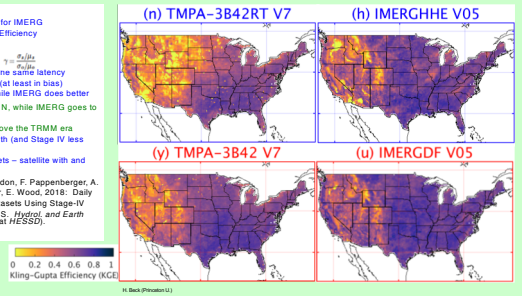
The time-series correlations for the IMERG runs are similar (below)
- Late improves on Early almost everywhere
  - this seems to be without regard to degree of bias, across the range of correlations
    - **this is an important result for data users**
  - but a few pockets of low correlation are resistant to change
    - does this tell us something about MRMS?
  - and Late still has some fairly fine-scale spatial structure
    - does this tell us something about MRMS?
    - does this tell us something about strangeness in the input data?



V05 VALIDATION – CONUS

Daily evaluation against Stage IV
- 2008-2017 for TMPA, 2014-2017 for IMERG
- Evaluated using the Kling-Gupta Efficiency

$KGE = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2}$,
where r = Pearson correlation, $\beta = \frac{\mu_e}{\mu_o}$ and $\gamma = \frac{\sigma_e/\mu_e}{\sigma_o/\mu_o}$
- IMERG improves over TMPA for the same latency
- In both, monthly gauge is helpful (at least in bias)
- TMPA falters north of ~40° N, while IMERG does better
  - TMPA calibration stops at 40° N, while IMERG goes to 65° N
  - the challenge in V06 is to improve the TRMM era
- The mountains are an issue in both (and Stage IV less sure)
- Statistics are shown for 26 datasets – satellite with and without gauge, and reanalyses:
  Beck, H., M. Pan, T. Roy, G. Weedon, F. Pappenberger, A. van Dijk, G.J. Huffman, R.F. Adler, E. Wood, 2018: Daily Evaluation of 26 Precipitation Datasets Using Stage-IV Gauge-Radar Data for the CONUS. *Hydrol. and Earth Sys. Sci.*, submitted (and posted at *HESSD*).



VERSION 06 UPGRADES

Morphing vector source switched to MERRA-2/GEOS-5 – see Tan poster

Morphed precip for all non-icy/snowy surfaces, including in polar regions

Full intercalibration to 2BCMB – V05 took shortcuts

Quality Index modified for half-hourly – see below

Modifications for TRMM era – primarily estimating the calibration for the band 35°-65° in both hemispheres

Revisions to internals raises the maximum precip rate from 50 to 200 mm/hr and no longer discrete

QUALITY INDEX – REVISED IN V06

Half-hourly QI (revised)
- approx. Kalman Filter correlation
  - based on times to 2 nearest PMWs
  - IR at time (when used)

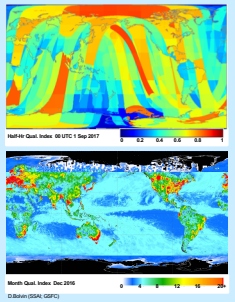
$$QI_h = \tanh\left(\sqrt{\sum \arctanh^2(r_i)}\right)$$

  - where r is correlation, and the i's are for forward propagation, backward propagation, and IR
  - approximate r when a PMW is used for just that satellite
- revised to 0.1° grid (0.25° in V05)
- thin strips due to inter-swath gaps
- blocks due to regional variations

Monthly QI (unchanged)
- Equivalent Gauge (Huffman et al. 1997) in gauges / 2.5° x 2.5°

$$QI_m = (S + r) * H * (1 + 10 * r^2) / e^2$$

  - where r is precip rate, e is random error, and H and S are source-specific error constants
- invert random error equation
- largely tames the non-linearity due to rain amount
- some residual issues at high values



VERSION 07 CONCEPTS

Multi-satellite issues
- Improve error estimation
- Develop additional data sets based on observation-model combinations
- Work toward a cloud development component in the morphing system

General precipitation algorithmic issues
- Introduce alternative/additional satellites at high latitudes (TOVS, AIRS, etc.)
- Evaluate ancillary data sources and algorithm for Prob. of Liq. Precip. Phase
- Track quality of PMW retrievals over snow/ice
- Work toward improved wind-loss correction to gauge data

SCHEDULE AND FINAL REMARKS

Early January 2019: begin Version 06 IMERG Initial Processing and Retrospective Processing
- The GPM era will be launched first, Final Run first
  - Early and Late retrospective processing use Final intermediate files, so they come after Final
  - complete data should take about a month
    - except Final is always ~3.5 months behind, so the Early and Late retrospective processing have to wait on Final Initial Processing to fill in the last 3 months of 2018
- The TRMM era will be launched after the GPM era is underway
- the Final-then-Early/Late pacing is true here as well
- complete data will take about 4 months using serial processing
- 4 km merged global IR data files continue to be delayed for January 1998-January 2000
  - the run will build up the requisite 3 months of calibration data starting from February 2000
  - the first month of data will be for June 2000
  - the initial 29 months of data will be incorporated when feasible

~2 years later: Version 07